4-23-2021

Seasonal Trend and Changepoint Analysis of United States Extreme Precipitation

Thea Sukianto
Boise State University

Meghan Edgerton
Boise State University

Mintaek Lee
Boise State University

Seasonal Trend and Changepoint Analysis of United States Extreme Precipitation

Abstract
We estimate seasonal trends in the contiguous United States extreme precipitation.

Seasonal Trend and Changepoint Analysis of United States Extreme Precipitation

Thea Sukianto, Meghan Edgerton, Mintaek Lee (Faculty Advisor: Jaechoul Lee)
Department of Mathematics, Boise State University

INTRODUCTION

In recent years, instances of extreme weather have increased in both frequency and severity in comparison to years past, from record high and low temperatures to flooding and landslides. These phenomena are often a byproduct of extreme precipitation. While many studies in the past have been done on average precipitation, extremes are of particular concern because of their significant consequences on both the physical and financial well-being of humans.

We aim to examine seasonal trends in extreme precipitation in the contiguous United States. The goal is to rigourously compute seasonal trends in United State seasonal maximum precipitation by considering undocumented changepoints at which the mean of the data changes. The changepoints are estimated using a genetic algorithm-based changepoint technique.

DATA

The data for this study was cultivated out of the United States Historical Climatology Network (USHCN), a detailed collaboration of 1,218 stations across the contiguous United States. The USHCN daily total precipitation recordings were aggregated to compute seasonal (spring, summer, fall, and winter) maximum precipitation. Our study only includes stations whose records spanned for at least 75 years and had less than 30% of the data missing or for 45-74 years with less than 7.5% missing rates. This criteria resulted in the use of 1,052 stations. Figure 1 displays spatial location of the 1,052 stations used in our seasonal extreme precipitation analysis.

Figure 1: Location of the selected 1,052 USHCN stations

METHODS

Extreme Value Model for Seasonal Maximum Precipitation

In analyzing extreme values of precipitation, a generalized extreme value (GEV) distribution was used to block seasonal maximum values. Since each season has varying climate dynamics, it is required that the model takes into account these differences. To model these features, we assume that each seasonal maximum follows a nonstationary GEV(μ_t,σ_t,ξ) distribution with parameters varying over seasons as follows:

$$\mu_{iT+\nu} = \beta_{0,\nu} + \beta_{1,\nu}\frac{iT+\nu}{100T} + \delta_{iT+\nu},$$
$$\ln(\sigma_{iT+\nu}) = \lambda_{0,\nu} + \lambda_1\frac{iT+\nu}{100T}.$$

The seasonal period $T = 4$ is set, with ν, a season indicator (ν = 1, spring; ν = 2, summer; ν = 3, fall; ν = 4, winter).

If a changepoint configuration is known, the parameters for the GEV distribution to be estimated are as follows:

$$\boldsymbol{\theta} = (\boldsymbol{\beta}^{\mathrm{T}}, \boldsymbol{\delta}^{\mathrm{T}}, \boldsymbol{\lambda}^{\mathrm{T}}, \xi)^{\mathrm{T}},$$

where the location, changepoint-induced mean shifts, and scale parameters of the GEV are defined as

$$\boldsymbol{\beta} = (\beta_{0,1}, \ldots, \beta_{0,4}, \beta_{1,1}, \ldots, \beta_{1,4})^{\mathrm{T}},$$
$$\boldsymbol{\delta} = (\Delta_1, \ldots, \Delta_c)^{\mathrm{T}},$$
$$\boldsymbol{\lambda} = (\lambda_{0,1}, \ldots, \lambda_{0,4}, \lambda_1)^{\mathrm{T}}.$$

These parameters are estimated by the maximum likelihood method, with the log-likelihood function for the parameter function $\boldsymbol{\theta}$ is determined using the GEV distribution function.

Changepoint Detection Via a Genetic Algorithm

A genetic algorithm (GA) is used to detect any significant mean shifts due to changepoints that were found in the seasonal precipitation maximums. As changepoints can significantly impact trend estimation for future time periods, detecting and analyzing them is an important key to accuracy in this study. To begin, an initial generation of *L=200* chromosomes is generated. Then, the GA procedure consists of the following steps:

1. Parent selection: Two chromosomes are selected.
2. Offspring production: A ‘child’ chromosome is produced in a probabilistic manner.
3. Offspring mutation: A mutation is allowed to prevent premature convergence to poor solutions (local minimums).
4. Steps 1-3 are repeated until convergence, and the solution is taken as the fittest chromosome in the terminating generation.

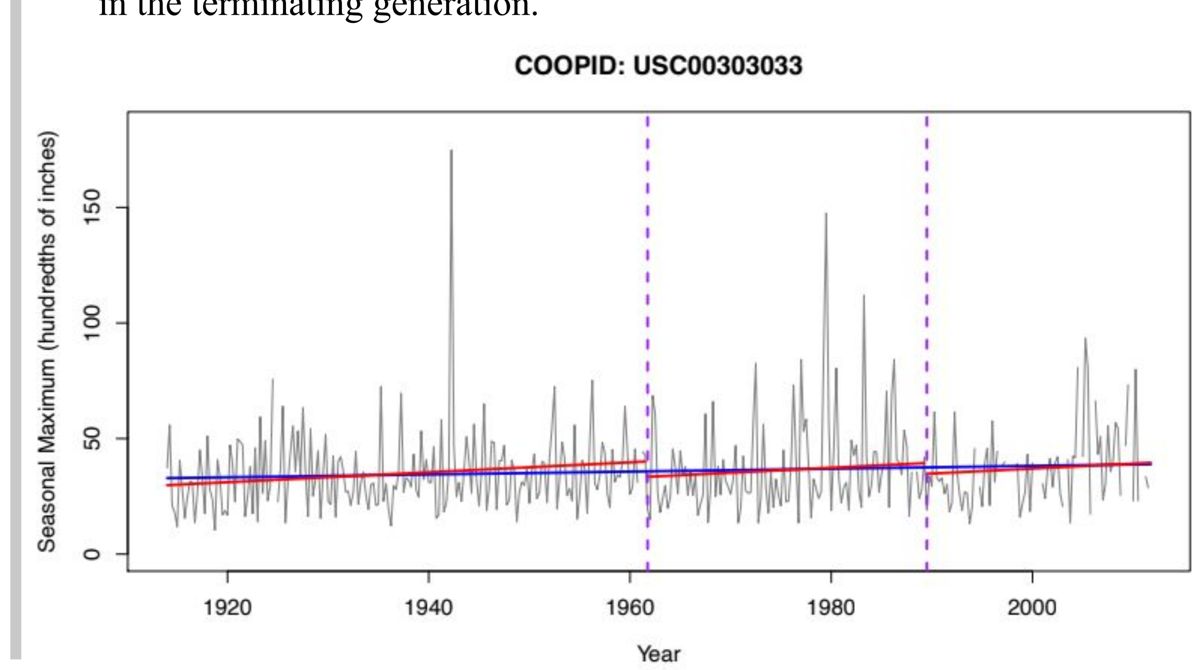

Figure 2: A plot generated from the GA of a station in Fredonia, NY showing estimated maximum precipitation trends (red and blue lines) and detected changepoints (purple dashed lines)

RESULTS

Seasonal Trend Maps

After running the GEV and the GA on all stations, seasonal trend maps (Figures 3-6) were generated considering the changepoints detected from the GA using inverse distance weighting (IDW) interpolation. IDW is a spatial interpolation method based on weighted averages such that the weights are the Euclidean distance between the interpolation location and a given station location raised to a power *p*. A leave-one-out cross-validation scheme is used to determine optimal values of *p*. For the sake of consistency, we choose *p* = 4 for all four maps.

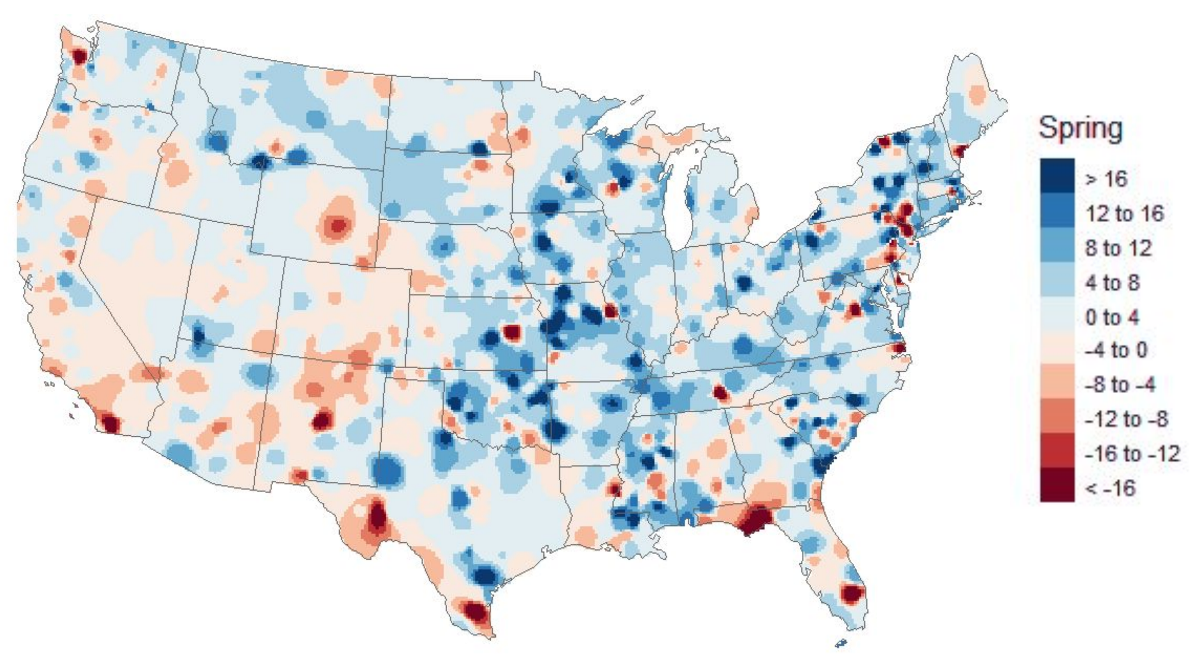

Figure 3: Map of estimated spring trends (in mm per century). Blue represents a positive trend and red a negative trend.

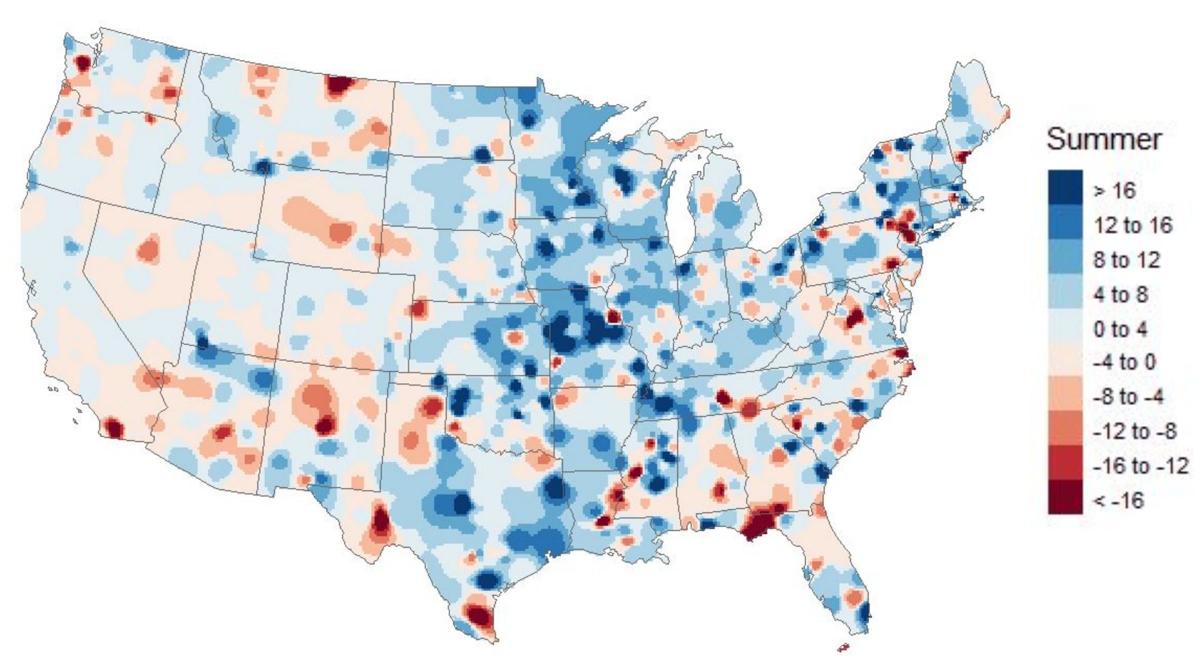

Figure 4: Map of estimated summer trends (in mm per century).

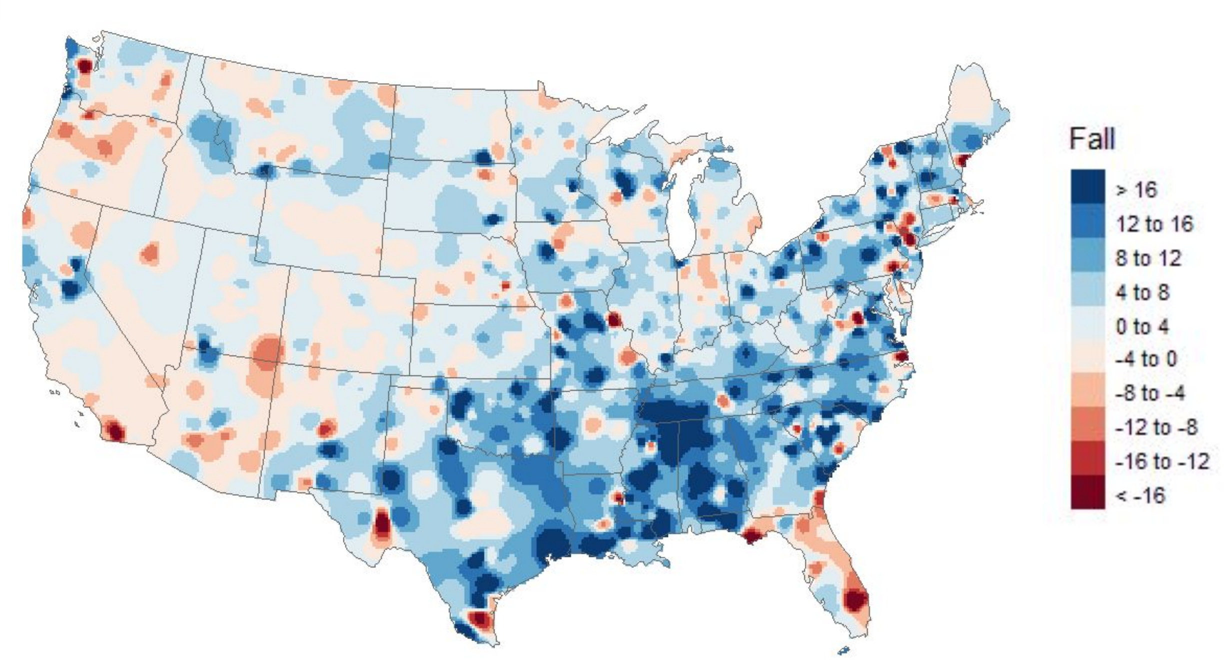

Figure 5: Map of estimated fall trends (in mm per century).

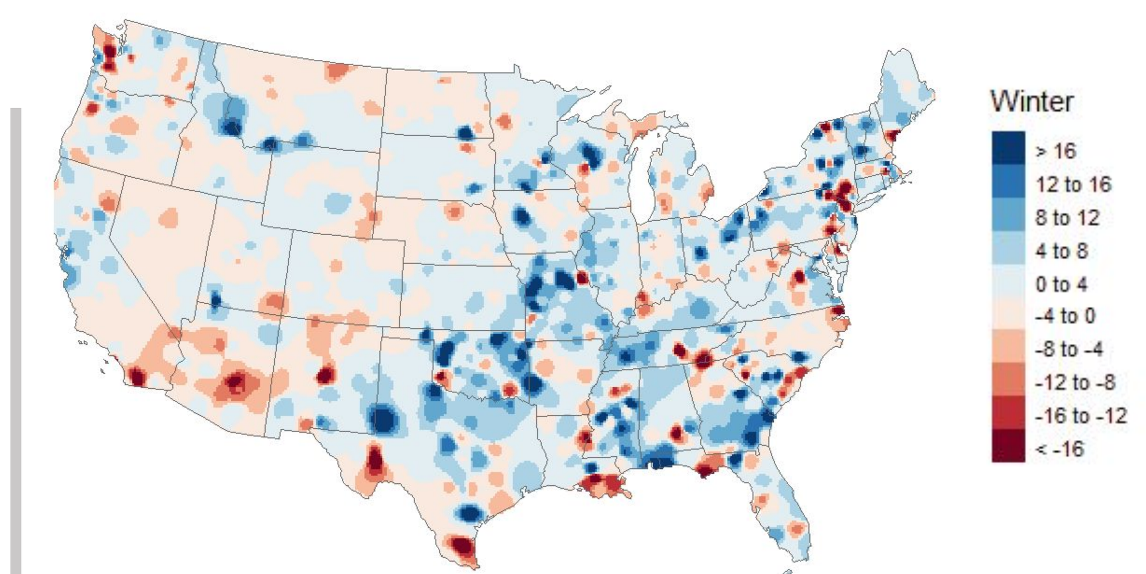

Figure 6: Map of estimated winter trends (in mm per century).

CONCLUSION

The seasonal trend maps produced from this study show more increasing maximum precipitation trends overall in the East, with the highest occurrence of increasing trends showing in the Fall season. These increasing trends in the Fall season lie predominantly in the Southern and Eastern regions of the United States. However, the Southwest regions show more decreasing trends for all seasons, apparently more severe in the Winter season. Following these results, further analysis will be conducted on selected candidate stations with significant changepoints in extreme precipitation. Considering these significant changepoints could be useful in preparation for weather that can seriously impact the physical and financial well-being of people in these regions.

REFERENCES

- Lee, M. and Lee, J. (2021). Long-term trend analysis of extreme coastal sea levels with changepoint detection. *Journal of the Royal Statistical Society: Series C (Applied Statistics),* 70, 434-458.
- Lee, J., Li, S., and Lund, R. (2014) Trends in extreme U.S. temperatures. *J. Climate*, 27, 4209–4225.
- Li, S. and Lund, R. (2012). Multiple changepoint detection via genetic algorithms. *J. Climate, 25*, 674-686.
- Lu, Q., Lund, R., and Lee, T. C. (2010). An MDL approach to the climate segmentation problem. *The Annals of Applied Statistics,* 4, 299-319.
- Lund, R. and Reeves, J. (2002). Detection of undocumented changepoints: A revision of the two-phase regression model. *J. Climate, 15*, 2547-2554.
- NCEI (2021). Billion-dollar weather and climate disasters, http://www.ncdc.noaa.gov/billions.
- Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 23rd ACM National Conference*. doi:10.1145/800186.810616.

ACKNOWLEDGEMENTS

The authors would like to thank the Boise State University, as well as our faculty mentor Jaechoul Lee from the Department of Mathematics at Boise State University for his valuable leadership in this project. Additionally, we would like to thank the United States Historical Climatology Network (USHCN) for making available the precipitation data used in this study.

(*Figure 3*): Map of estimated spring trends.